The Use of Decision Tree and Multiscale Texture for Classification of JERS-1 SAR Data over Tropical Forest

Marc Simard, Sasan S. Saatchi, and Gianfranco De Grandi, *Senior Member, IEEE*

***Abstract*—The objective of this paper is to study the use of a decision tree classifier and multiscale texture measures to extract thematic information on the tropical vegetation cover from the Global Rain Forest Mapping (GRFM) JERS-1 SAR mosaics. We focus our study on a coastal region of Gabon, which has a variety of land cover types common to most tropical regions. A decision tree classifier does not assume a particular probability density distribution of the input data, and is thus well adapted for SAR image classification. A total of seven features, including wavelet-based multiscale texture measures (at scales of 200, 400, and 800 m) and multiscale multitemporal amplitude data (two dates at scales 100 and 400 m), are used to discriminate the land cover classes of interest. Among these layers, the best features for separating classes are found by constructing exploratory decision trees from various feature combinations. The decision tree structure stability is then investigated by interchanging the role of the training samples for decision tree growth and testing. We show that the construction of exploratory decision trees can improve the classification results. The analysis also proves that the radar backscatter amplitude is important for separating basic land cover categories such as savannas, forests, and flooded vegetation. Texture is found to be useful for refining flooded vegetation classes. Temporal information from SAR images of two different dates is explicitly used in the decision tree structure to identify swamps and temporarily flooded vegetation.**

***Index Terms*—Decision tree, forest, multiscale, synthetic aperture radar (SAR).**

I. Introduction

In RECENT years, there has been a growing interest in the use of synthetic aperture radar (SAR) data in remote sensing for applications ranging from land cover mapping to change detection. In particular, remote sensing using SAR data has the potential of becoming the most practical method for mapping and monitoring land cover over the tropics, where continuous cloud cover hinders optical imagery [1]–[6]. For these applications, it is important to develop tools to obtain useful thematic information from radar data in terms of landscape features and patterns. In this paper, we use a decision rule approach for classification of SAR data that exploits multiscale texture features of SAR imagery.

Manuscript received September 6, 1999; revised April 29, 2000. This work was supported by a grant from the National Aeronautics and Space Administration (NASA), Washington, DC.

M. Simard is with the Jet Propulsion Laboratory, Pasadena, CA 91109 USA (e-mail: simard@innu.jpl.nasa.gov).

S. Saatchi is with the Radar Science and Engineering Section, Jet Propulsion Laboratory, Pasadena, CA 91109 USA.

G. De Grandi is with the Space Application Institute, Joint Research Centre, Ispra (VA), Italy (e-mail: frank.de-grandi@jrc.it).

Publisher Item Identifier S 0196-2892(00)08902-6.

The backscatter spatial statistics are given for a specific scale and depend on the distribution of scattering elements on the imaged surface. Multiscale texture fully exploits the information carried by scale variations in SAR imagery. It is known that texture measures derived from SAR data often do not follow a normal distribution [7], [8]. As a result, classification approaches such as Bayesian maximum likelihood estimation (MLE), which assume a normal distribution of data, are not suitable for SAR image classification based on a feature space comprising texture measures. Theoretically, techniques that do not rely on the statistical distribution of the image data are more suitable for SAR multiscale classification. Decision tree classifiers are among those techniques and have been successfully applied to remote sensing data [9]–[12].

We demonstrate the application of the decision tree classifier with multiscale texture features in a test case based on Japanese Earth Resource Satellite (JERS-1) L-band SAR data acquired over tropical forests. In particular, we will assess the classification performance and investigate feature contributions and feature relations as inferred from the decision tree structure.

The present investigation is carried out in the context of the Global Rain Forest Mapping Project (GRFM). The project was initiated in late 1995 by the National Agency for Space Development of Japan (NASDA), Tokyo, Japan. Postprocessing and compilation of the GRFM Central Africa mosaic was performed at the European Commission Joint Research Center, Ispra, Italy, within a collaboration agreement with NASDA [13]–[15]. The 100 m resolution GRFM data sets are currently available to interested users. One of the main objectives of GRFM is to assess the potential of JERS-1 SAR data for mapping the vegetation cover and monitoring land cover changes in the tropics. Thematic maps derived from JERS-1 data over Central Africa will be extremely valuable given the difficulty in this region of getting reliable and spatially explicit data or survey information on the status of vegetation cover.

In this paper, we focus on the classification methodology and analyze the performance of the classifier over a small area in coastal Gabon. In Section II, we describe the study area and the characteristics of remote sensing data. In Section III, the classifier is presented with a brief introduction to the multiscale texture maps. In Section IV, we discuss the training data set and the application of classifier on the image data. In Section V, the contribution of the features and the stability of the classifier are examined using the decision tree structure. Finally, we analyze and verify the results of the classifier.

II. Data and Study Area

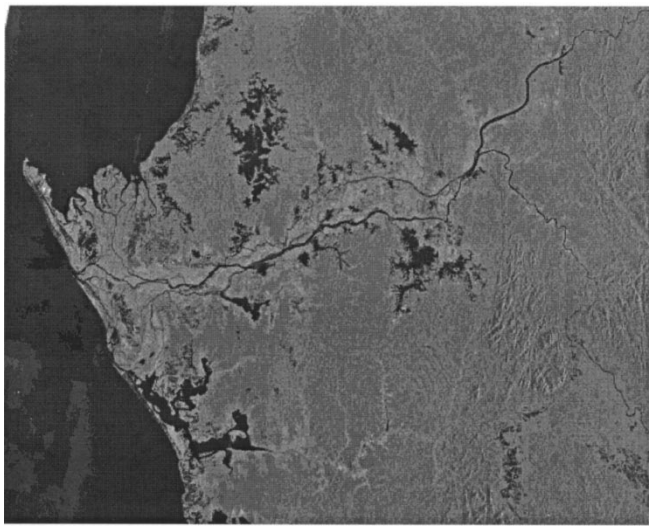
We use a reduced data set extracted from two GRFM mosaics of JERS-1 SAR images covering the entire Central Africa [16], [13]. The acquisition dates of the images, February and November 1996, correspond to the high and low water seasons of the Congo River respectively. The GRFM Africa mosaics are derived from Level 2.1 NASDA high resolution and georeferenced JERS-1 products (3-look, 12.5 m pixel spacing in both azimuth and range). In this study, we use the Africa mosaic baseline product with 100 meter resolution [13], [8]. The 100 m amplitude backscatter images were generated by low-pass filtering and decimation from the high resolution NASDA products. The downsampling process is based on a multiresolution decomposition where the scaling function is a cubic spline polynomial.

All data is reduced to bytes for storage and processing purposes. We used amplitude (square-root of intensity) instead of intensity because: 1) the histogram conserves a larger dynamic range of bytes for medium backscatter targets, such as forest, which are the most interesting targets in out thematic context; and 2) intensity enhances high backscatter targets (e.g. urban areas and floods), while a logarithm representation enhances low backscatter targets (e.g. calm waters and savannas).

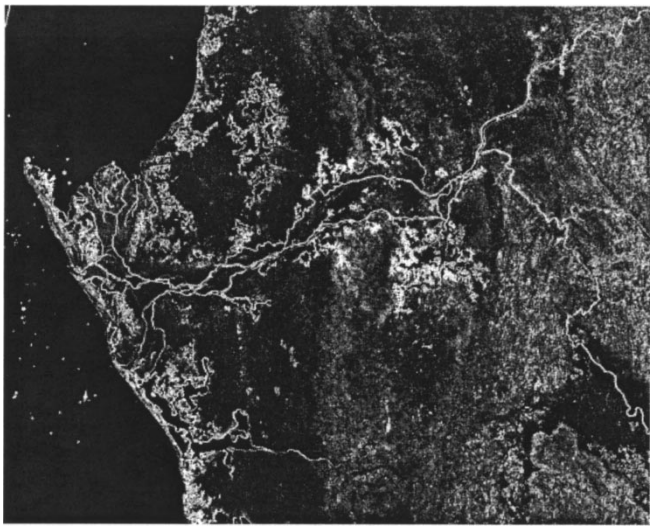
The study area covers a region on the western coast of Gabon in Central Africa (Fig. 1). The region extends from 0°5’ 19” S to 1° 53’ 50” S latitude and 8°E to 10° 52’ 13” E longitude and includes a major portion of the Ogooué River basin along the coast of Gabon. The Ogooué River originates in the mountain ranges of northern Lékoumou in Congo and crosses a variety of landscapes from grassland and woody savannas to evergreen dense forests. Along the coast, it creates a basin with a variety of floodplain vegetation types and landscapes suitable for fisheries and agriculture. The eastern portion of the study area has gently sloped hills. Along the Ogooué basin in the west, the landscape is flat, resulting in areas of very slow drainage and creating small wet and sometimes inundated pockets along the river and coast.

In order to identify training classes for the classifier, we used land cover maps provided by the Ministère des eaux et forêts et du reboisement (LCMG) (Ministry of Water and Forests and Reforestation). We also used a variety of field data collected by us and our collaborators during field experiments carried out in 1997 (personal communication with Chris Wilks and Lee White). The main vegetation categories in the region are as follows.

- Forest: The forest category is divided into two classes: open secondary and old secondary forests. The open secondary forests have a relatively low backscatter in the image (with respect to old secondary). These forests are more than ten years old with fast growing tree species that often create a low and open forest canopy (over 50% of the canopy area is open to light). In the classification procedure we refer to this class as “Open forest” (*Open*). On the other hand, old secondary forest, which is the dominant type in the region, has higher radar backscatter values than open secondary and appears with different textures. These forests have a closed canopy. A portion of the old secondary forests in the region contain the Okoume trees, which are often logged. There is primary forest in the region often found near the mountains in the east. However, primary forests are very similar to old secondary forest both in phenological terms and in the radar image. Thus, the old secondary and primary forest are combined in the “forest” class (*For*).
- Savannas: We separate the savannas into two classes: “Grass Savanna” (*Sav*) and “Woody Savanna” (*Woo*). The latter is not identified on the available land cover map. However, the existence of this class is documented by ground surveys. The Woody Savanna has seasonal characteristics and can also host extensive fire events. These facts are reflected in relative changes between the backscatter values of the two acquisition dates.
- Floodplains: We chose five land cover types for floodplains: “Permanently” (*Perm*) flooded and “Temporarily” flooded woody vegetation (*Tmp*), and “Low Mangroves” (*Mang*), “Swamp vegetation” (*Sw*), and “Raphia” (Palmaceae) (*Raph*). The flooded woody vegetation consists



(a)



(b)

Fig. 1. (a) Amplitude SAR image of the study site in Gabon. The image is an extract of the GRFM mosaic with 100 m resolution. (b) Wavelet-based texture map of the region at scale 400 m.

of forests that are inundated during both acquisition dates or only in the high water image. The Low Mangroves are located along the coast and inland along the river, and are primarily Black Mangroves (Avicennia germinans). The plants are short and dense and usually cover the region like a carpet along the intertidal zones. "Swamp vegetation" (*Sw*) refers to a variety of floodplain vegetation ranging from low grass and shrubs to low density woody plants that are waterlogged for most of the year. "Raphia" (Palmaceae) (*Raph*) forests are scattered primarily along lake edges and river basins far from the coast.

- Urban: The "Urban" class (*Urb*) is easily distinguished on the SAR images due to the presence of corner reflectors. The two major urban areas in the image are the city of Port Gentil, at the Baje of Cap Lopez and the city of Lambaréné along the Ogooué River.

III. Classification Method

A. The Decision Tree Algorithm

The decision tree algorithm used in this paper is based on the algorithm described in [17]. The decision tree classifier is a set of hierarchical rules which are successively applied to the input data. Those rules are thresholds used to split the data into two groups (binary splits). Each split (also called node) is such that the descendant nodes contain more homogeneous data samples (i.e., the nodes are purer in terms of classes). Many layers (features) can be input into the decision tree to refine class description. A split is chosen because of its ability to render the nodes purer based on a purity measure $i(t)$ and can be determined by any single feature.

Decision tree classifiers have been successfully applied to remote sensing data in the past [10], [11], [18]. As the decision tree approach does not rely on any *a priori* statistical assumption, it is suitable for classifying SAR image data. Moreover, decision tree rules are explicit and allow for identification of features which are relevant to distinguish specific classes. Thus, the analysis is reduced to the most useful layers. The structure of the decision tree can also reveal hierarchical and nonlinear relationships among input layers. These relationships often result in a given class being described by various terminal nodes. Terminal nodes are the final decisions, which assign a sample x to class j. The variable x is a feature vector describing one sample from the input population.

The decision tree algorithm is supervised because it relies on training samples to grow. A set $S_g(x,\ j)$ of independent variable x with a dependent variable j (the label or class) is input into the decision tree growing algorithm. The split is chosen to maximize the reduction of impurity ($\Delta i(t)$) (or cost) of the parent node t with respect to the impurity of its child nodes such that

$$\Delta i(t) = i(t) - p_l i(t_l) - p_h i(t_h). \tag{1}$$

Variables p_l and p_h are the total proportions of samples in node t that reach into child nodes t_l and t_h. The optimization of the impurity is performed by Gini criterion. This criterion was shown to be generally efficient and can also be shown to minimize the resubstitution estimate $R(T)$ [defined in (4)] for the minimum square error [17, p. 124]. The criterion is defined as

$$1 - \sum_j p^2(j|t). \tag{2}$$

It uses a rule that assigns a sample x randomly selected from node t to class j with probability $p(j|t)$. This probability is

$$p(j|t) = \frac{P_J \cdot N_j(t)/N_j}{\sum_j P_J N_j(t)/N_j} \tag{3}$$

where P_J is the *a priori* probability for class j, and N_j and $N_j(t)$ are the number of class j samples in $S_g(x,\ j)$ and node t, respectively. In the absence of any *a priori* information about the class distribution, P_J is chosen to be equal for all classes. However, in principle, P_J could be used as a class weight to favor classification for given classes. In our case, all classes are equally important. Note that the sum over j in (2) allows for a global decision that emphasizes the purity of the groups rather than the purity of an individual class. In theory, the decision tree can grow indefinitely until all classes are separated. This process can lead to a very large decision tree (eventually resulting in one sample in each terminal node). To circumvent this problem, we will grow a manageable size decision tree (nine successive splits or levels), which results in 2^8 terminal nodes. The size of this initial decision tree is further reduced by the "pruning" process.

The pruning process consists in eliminating the inefficient (or weak) branches of the decision tree. The pruning is based on the cost-complexity function $C_\alpha(T_t)$ of nonterminal node t [17]. The function is defined as $C_\alpha(T_t) = R(T_t) + \alpha \cdot nT$, with $R(T_t)$, the resubstitution estimate of the set of terminal nodes of subtree T_{ts} under t. $C_\alpha(T_t)$ takes into account the number of terminal nodes nT under t, which characterizes the size of T_{ts}. The variable α may vary continuously. However, a decision tree is finite, and the decision tree $T(\alpha)$ that minimizes $C_\alpha(T)$ is optimum until some value α is reached. Then, $C_\alpha(T_t)$ is as large as $C_\alpha(t) = R(t) + \alpha$. That is, the cost complexity of node t itself. At this point, only node t is kept as a terminal node since its children are of the same complexity. The weakest nonterminal node is the one with minimum

$$\alpha(t) = \frac{R(t) - R(T_t)}{nT - 1}$$

with resubstitution estimate

$$R(t) = r(t)p(t),$$

and

$$r(t) = 1 - \max_j p(j|t) \tag{4}$$

where $p(t)$ is the probability that any sample fall in node t. The weakest node t is one with many terminal nodes nT and with a low decrease in impurity with respect to terminal nodes of T_t. Pruning of the initial decision tree T is illustrated in Fig. 2. It shows how the weakest nodes of the initial decision trees are iteratively cut ($T - T_{ts}$) to produce a sequence of subtrees T_k with different complexities. This sequence is obtained by saving

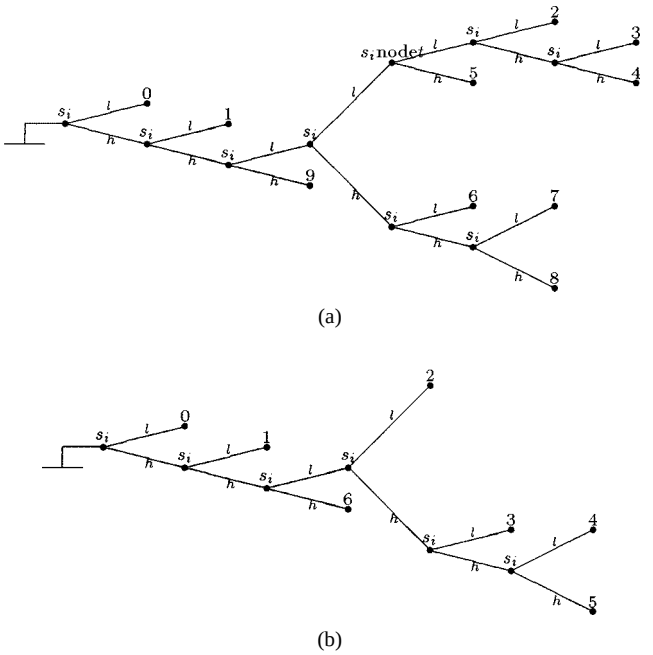


Fig. 2. (a) Example of a decision tree T. The set of terminal nodes under node t are $T_t = [2, 3, 4, 5]$. (b) Decision tree $T - T_{ts}$ resulting from the pruning of decision tree T of its subtree T_{ts}, which contains all nodes T_t. At each node, a value correponding to the best split s on the input feature i is given (s_i). The branches are denoted by l and h, which correspond sample feature values lower and higher than thresholds s, respectively.

each decision tree ($T - T_{ts}$). The optimum-size decision tree is chosen using an independent testing set $S_p(x, j)$ of samples of known classes j. The independent samples are input into every decision tree of sequence T_k and estimates of $R_m(T_k)$ are calculated as

$$R_m(T_k) = \sum_{i,j} N_{i,j}/N_j \tag{5}$$

which is the proportion of class j samples classified as class i using decision tree T_k. The function $R_m(T_k)$ measures the overall misclassification rate of decision tree T_k.

The development of the decision tree is influenced by noise, which causes intraclass variability in the data. We consider the noise in the data as speckle noise and random distribution of forest radar scatterers. The effect of noise on decision tree structure is discussed in [17, p. 155]. Because of noise, sample x is randomly distributed within a class. The main effect is to randomize the choice of the feature on which to apply a split. This is caused by a similar decrease in impurity Δi for competing features. We expect that the larger the noise, the tighter the competition between splits. This competition will mainly be between redundant (correlated) features. As discussed in [8], at large scales, the contribution of speckle noise to multiscale image texture is lower than the contribution of forest structures contribution. For example, in the GRFM 100 m SAR data over tropical forest, we expect the forest large scale structures to be the main cause of intraclass variability. Note that the GRFM mosaics have 59 equivalent number of looks [13]. However, in the high resolution 3-look SAR data, such as the 12.5 m pixel amplitude (square root of intensity) images, the multiplicative noise (speckle) will contribute significantly [19]. As a result, the intraclass variability is dependent on the class mean backscatter. In this case, we would expect more split competition and branches for the high backscatter regions. Speckle filtering would probably improve results in the high resolution case.

We added two minor tests to the basic binary decision tree testing algorithm due to overspecialization of the nodes. Two cases can occur, mainly in late nodes.

1) A node t_e is empty. While using an independent sample set $S_p(x, j)$ for testing sequence T_k, some nodes t_e, created using $S_g(x, j)$, might be empty ($N_j(t_e) = 0$) of samples x of $S_p(x, j)$.
2) A terminal node t_c changes class label. This change occurs when the class memberships $N_j(t')/N_j$ within a terminal node t' are similar. The initial class label j of t', computed from the growth training set $S_g(x, j)$, might change to j' when using testing set $S_p(x, j)$. That is, $\max(N_j(t_c)/N_j)$ for $S_g(x, j) \neq \max(N_j(t_c)/N_j)$ for $S_p(x, j)$.

In the first case, we estimate the contribution of the father split t_{fe} of t_e. The misclassification $R_m(T_{ts_{fe}})$ of the subtree $T_{ts_{fe}}$ under t_{fe} is computed using $S_g(x, j)$ samples. Note that t_{fe} does not contribute to classification of $S_p(x, j)$. Thus, T_k could be pruned from t_{fe} to reduce further complexity. The decision tree T_k containing node t_e is pruned only if t_{fe} does not improve significantly the classification accuracy. We arbitrarily chose a minimum 2% accuracy improvement from t_{fe} alone to terminal nodes of $T_{ts_{fe}}$.

In the second case, we use samples of $S_g(x, j)$ and $S_p(x, j)$ to re-evaluate the misclassification $R_m(T_{ts_{fc}})$ of the subtree $T_{ts_{fc}}$ under t_{fc}. The class label j that has the largest total membership $\sum_{S_p(x), S_g(x)}(N_j(t_c)/N_j)$ in terminal node t_c is assigned to t_c. Again, a 2% minimum classification accuracy improvement is required. This correction might allow a split resulting in one noisy and one pure terminal node instead of pruning up to the noisy parent node. Finally, the optimum decision tree is one with the minimum estimate $R_m(T_k)$.

B. Multiscale Texture

In addition to the two season radar data, texture images were used as input to the classifier. The texture images are derived from radar images by using Mallat's discrete wavelet transform (DWT) algorithm [20] without decimation.

DWT has been successfully applied as a texture descriptor for classification of image data [21]–[24]. The DWT produces a low resolution approximation of the original image as well as 3 detail images, one for each orientation. Consequently, a texture map is generated by combining the three orientation detail images using quadratic addition and normalizing by the low resolution approximation. These texture maps are then generated at 200, 400, and 800 m by iterating on the 100 m data and the low resolution approximations at each scale. At each scale, the texture maps provide information about the surface heterogeneity and enhance structures such as edges [25]. For more information on texture maps, the reader is referred to [8].

TABLE I
CLASS DEFINITIONS: THIS TABLE DESCRIBES THE CLASS CODES, THE SIZE N OF THE GROWTH TRAINING SAMPLE S_g, TEST SAMPLE S_p, AND VALIDATION SAMPLE S_v WITH THE DESCRIPTION OF CLASSES. N.B. LCMG IS THE LANDCOVER MAP OF THE "MINISTÈRE DES EAUX ET FORÊTS ET DU REBOISEMENT" OF GABON

class	code	N S_g	N S_p	N S_v	Class Description
Low Mangroves	0 Mang	96	86	27	Contains low mangroves. Defined as low-mangrove on LCMG and bright on SAR image with a filamentary structure.
Urban	1 Urb	60	42	20	Defined as urban on LCMG and very bright on SAR image
Swamps	2 Sw	160	176	45	Defined as flooded low vegetation on LCMG and characterized by temporal change on the SAR image. It is bright on the low water SAR image and dark on high water SAR image.
Temporarly flooded	3 Tmp	111	60	73	Defined as temporarly flooded on LCMG. Characterized by temporal change on the SAR image. It is brighter in the high water SAR image.
Permanently flooded	4 Perm	104	63	37	Defined as permanently flooded on LCMG. Characterized by no temporal change. It is bright on both SAR images.
Woody savanna	5 Woo	72	42	52	Defined as savannas on LCMG. It is bright on one SAR image.
Forest	6 For	1386	645	423	Defined as old secondary and primary forest on LCMG. Medium tones.
Grass Sav.	7 Sav	1038	540	192	Defined as savannas on LCMG. Dark on both SAR images.
Open forest	8 open	210	105	231	Coincide with young secondary on LCMG. Contains misclassified high mangroves. Characterized by medium low backscatter.
Raphia	9 Raph	702	234	136	Defined as Raphia on LCMG. Characterized by high backscatter and smooth texture on SAR images.

IV. CLASSIFICATION PROCESS

Because the decision tree classifier is a supervised method, the choice of training data is critical in the process of classification. The training areas are based on the *a priori* knowledge of the region. In this test case, we have used the vegetation map of the coastal region of Gabon provided by "Ministère des eaux et forêts et du reboisement," and the notes collected during the 1997 field campaign in the region. In choosing the training polygons, we have also used our knowledge of the SAR response to various surface features. We include expert information because 1) the vegetation map is coarse in scale and has less detailed information than SAR data, and 2) the map is produced by visually interpreting aerial photography and optical data. This process introduces classification errors that are not quantified and therefore the map cannot be taken as ground truth with a given confidence. One major discrepancy between the vegetation map and the thematic interpretation of the SAR data is related to the Temporary and Permanently flooded vegetation along the coast. At the time of the SAR data acquisition, the inundation patterns were different from the ones documented in the vegetation map. Due to the sensitivity of the radar backscatter data to the ground water status, the inundation patterns captured by the SAR imagery are probably reliable. However to be on the safe side, because of its sensitivity to inundation [26], [27], we have used for training areas only those areas where both data sets are in agreement. Moreover, in collecting the training data sets, we avoided those areas that showed discontinuity in the SAR response but were labeled as one class in the vegetation map. This is primarily due to changes in land cover and land use during the time between 1989 (acquisition of the data for the vegetation map) and 1996 (the SAR data take).

Intraclass variability may also appear when choosing the training data set as a result of the spatial proximity of other classes. The use of low resolution features increases the number of mixed pixels. Indeed, low resolution approximations and texture measures are strongly influenced near edges. As a result, the near-edge region may be assigned to a wrong class. While this phenomenon is often neglected in classification, it becomes important when using relatively large analysis windows or multiscale analysis. To overcome mixed pixels due to edges (discontinuity), the training areas included pixels near and far from the edge. This allows the classifier to use high resolution images to correct for low resolution mixed pixels if required.

During the process of classification, we noticed that water and grassland savanna were not distinguished visually or by the decision tree classifier Fig. 1. Other studies have also pointed out this problem when classifying the L-band SAR data [28]. To avoid the misclassification, we used the Digital Chart of the World to mask the open water bodies in the image. The training areas were chosen over the ten land classes and are included in Table I. We chose three sets of data for training, pruning, and validation (S_g, S_p, S_v). The number of pixels in each set is given in Table I.

After the training set selection, the next step in the classification process is the organization of input layers. We generated a total of seven feature images to be used as input layers to the classifier as described in Section III. These include

1) low water 100 m data (GRFM mosaic data) (LWHR);
2) high water 100 m data (GRFM mosaic data) (HWHR);
3) low water 400 m low resolution approximation (LWLR);
4) high water 400 m low resolution approximation (HWLR);
5) texture at 200 m (TX200);
6) texture at 400 m (TX400);
7) texture at 800 m (TX800).

We used only the low water texture features, because they are highly correlated with the high water texture features. The initial decision tree was arbitrarily limited to a maximum of nine successive splits (or levels). This is equivalent to 2^8 terminal nodes and is sufficient to classify all ten classes. We first used

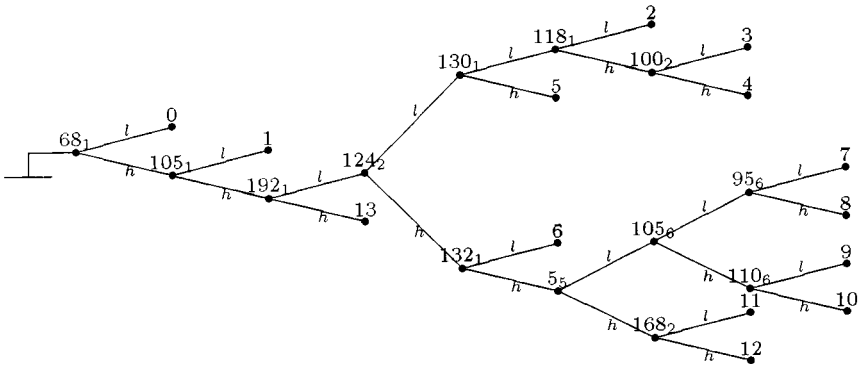


Fig. 3. Decision tree obtained from seven features. At each node, the best split s is given with the input feature i (s_i). The input features are the amplitudes for 100 m low water ($i = 0$), 400 m. (1) Low water, (2) 400 m high water, (3) 100 m high water, (4) textures at 200 m, (5) 400 m, and (6) 800 m. The branches are denoted by l and h, which are values lower and higher than threshold s, respectively. The class memberships associated with the terminal nodes of this decision tree, are tabulated in Table II.

the seven input features: LWHR, HWHR, LWLR, and HWLR and TX200, TX400, and TX800. The weakness of all nonterminal nodes was estimated using (4), and the sequence of pruned decision trees T_k was created. The sample set S_p is the independent sample to test sequence T_k. Finally, the optimum decision tree was chosen.

The optimum decision tree has 14 terminal nodes. These nodes represent all ten classes because various nodes may carry the same class label. The final decision tree for seven features is shown in Fig. 3. Each terminal node t' of the decision tree is then identified to the class with largest membership $N_j(t')/N_j$ (see Table II). In Table III, the rows represent terminal nodes and the columns the classes.

Before analyzing the results, we also built three other decision trees with various combinations of features. We therefore produce four classifications using the following feature combinations:

- seven features (LWHR, HWHR, LWLR, HWLR, TX200, TX400, TX800);
- five features (LWHR, HWHR, TX200, TX400, TX800);
- two features (LWHR, HWHR);
- four features (LWHR, HWHR, LWLR, HWLR).

We refer to these decision trees as T7, T5, T2, and T4, respectively. Each of these decision trees will help us understand the contribution of features and the decision tree sensitivity to input layers. The results are shown in terms of confusion matrix in Table IV. A comparison of the resulting classification maps is shown in Fig. 4.

TABLE II
CLASS MEMBERSHIPS (%) USING SEVEN FEATURES: MEMBERSHIP IS THE PROPORTION (PERCENT) OF CLASS j (COLUMN) SAMPLES OF BOTH TRAINING SETS $S_g(j)$ AND $S_p(j)$ THAT FALL INTO TERMINAL NODE T_n (ROW). THE FEATURES ARE 100 M RESOLUTION LOW AND HIGH WATER AMPLITUDE DATA, 400 M RESOLUTION LOW AND HIGH WATER AMPLITUDE DATA, WITH TEXTURE MAPS AT 200, 400, AND 800 M. THE ASSOCIATED DECISION TREE IS SHOWN IN FIG. 3

class / nodesn	0 Mang	1 Urb	2 Sw	3 Tmp	4 Perm	5 Woo	6 For	7 Sav	8 Open	9 Raph
0	0	0	0	0	0	10	0	100	0	0
1	0	0	3	0	0	90	0	0	15	0
2	0	0	3	14	0	0	11	0	85	0
3	0	0	18	0	0	0	0	0	0	0
4	1	0	2	1	0	0	87	0	1	0
5	2	0	73	0	3	0	1	0	0	0
6	1	0	0	77	0	1	1	0	0	0
7	0	0	0	2	7	0	0	0	0	0
8	0	0	0	1	10	0	0	0	0	3
9	0	0	0	1	0	0	0	0	0	5
10	0	0	0	2	0	0	0	0	0	77
11	88	4	0	4	51	0	0	0	0	14
12	8	6	0	0	28	0	0	0	0	0
13	0	91	0	0	1	0	0	0	0	0

V. ANALYSIS AND DISCUSSION

In this section, we discuss the impact of using different features on overall and individual classification accuracy. No consistent dependency was found between the number of terminal nodes versus the number of features ($nT_{T7} = 14$, $nT_{T5} = 25$, $nT_{T2} = 17$, $nT_{T4} = 20$). The fact that T7 is the smallest decision tree is due to the use of the "good" features as opposed to the inclusion of all features. As seen in Fig. 3, only features 1, 2, 5, and 6 are used (corresponding for T7 to LWLR, HWLR, TX400, and TX800, respectively). On the other hand, this does not imply that unused features are "bad" information. As discussed earlier, competition between splits can be very tight for a sample set S_g. In fact, a feature is best judged by what can not be achieved without its use. Comparison of the various decision trees can give insight into this issue.

The overall misclassification rates (commission error) of the decision trees are: T7: 16%, T5: 21%, T2: 34%, and for T4: 21%. The use of all features provided the best overall classification, because it allowed the decision tree algorithm to choose the best splits amongst a variety of features. However, some individual classes are best classified using fewer features. This is because an early split on one feature might mix classes while purifying others, in search of the best global solution.

All decision trees provide a high classification rate for the Urban class and Grass savannas, because of their distinct backscatter. However, performances differ significantly for other classes. This demonstrates the importance of feature choices in the feature space. In the remainder of the section, we discuss the contribution of the different features for classification of individual classes. We will also briefly discuss

TABLE III
CLASS MEMBERSHIP USING SEVEN FEATURES FOR EXPLORATORY DECISION TREE: MEMBERSHIP IS THE PROPORTION (PERCENT) OF CLASS j (COLUMNS) SAMPLES OF BOTH TRAINING SETS $S_g(j)$ AND $S_p(j)$ THAT FALL INTO TERMINAL NODE T_n (ROWS). THE FEATURES ARE 100 M RESOLUTION LOW AND HIGH WATER AMPLITUDE DATA, 400 M RESOLUTION LOW AND HIGH WATER AMPLITUDE DATA, WITH TEXTURE MAPS AT 200, 400, AND 800 M. THE ASSOCIATED EXPLORATORY DECISION TREE IS SHOWN IN FIG. 5

A)	Mang	Urb	Sw	Tmp	Perm	Woo	For	Sav	Open	Raph
Mang	88	4	0	4	51	0	0	0	0	0
Urb	0	91	0	0	1	0	0	0	0	0
Sw	2	0	91	0	3	0	1	0	0	0
Tmp	1	0	0	77	0	1	1	0	0	0
Perm	8	6	0	3	45	0	0	0	0	17
Woo	0	0	3	0	0	90	0	0	15	0
For	1	0	2	1	0	0	87	0	1	0
Sav	0	0	0	0	0	10	0	100	0	0
Open	0	0	3	14	0	0	11	0	85	0
Raph	0	0	0	3	0	0	0	0	0	82

B)	Mang	Urb	Sw	Tmp	Perm	Woo	For	Sav	Open	Raph
Mang	80	2	1	2	28	0	0	0	0	0
Urb	0	90	0	0	0	0	0	0	0	0
Sw	1	1	69	0	1	0	4	0	6	0
Tmp	1	0	1	63	0	7	4	0	10	3
Perm	15	6	0	14	69	0	0	0	0	6
Woo	0	0	1	1	0	91	0	0	5	0
For	0	0	6	4	1	0	72	0	7	3
Sav	0	0	0	0	0	0	0	100	0	0
Open	2	1	22	4	0	1	17	0	69	0
Raph	0	0	0	11	1	0	2	0	0	85

C)	Mang	Urb	Sw	Tmp	Perm	Woo	For	Sav	Open	Raph
Mang	57	6	4	1	51	0	4	0	2	25
Urb	0	90	0	0	0	0	0	0	0	0
Sw	0	1	90	0	0	0	7	0	8	0
Tmp	8	0	0	66	2	1	4	0	3	16
Perm	11	1	0	2	27	0	0	0	0	18
Woo	0	0	1	1	0	92	1	0	11	0
For	1	3	1	26	0	0	81	0	58	2
Sav	0	0	0	0	0	0	0	100	0	0
Open	0	0	5	0	0	8	2	0	19	0
Raph	21	1	0	5	20	0	0	0	0	39

D)	Mang	Urb	Sw	Tmp	Perm	Woo	For	Sav	Open	Raph
Mang	54	4	0	0	31	0	0	0	0	16
Urb	0	91	0	0	1	0	0	0	0	0
Sw	0	0	91	0	1	0	0	0	0	0
Tmp	1	0	0	87	0	1	3	0	4	0
Perm	11	5	0	0	35	0	0	0	0	4
Woo	0	0	3	0	0	90	0	0	15	0
For	1	0	2	1	0	0	87	0	1	0
Sav	0	0	0	0	0	10	0	100	0	0
Open	0	0	3	4	0	0	10	0	80	0
Raph	33	0	0	8	31	0	0	0	0	79

E)	Mang	Urb	Sw	Tmp	Perm	Woo	For	Sav	Open	Raph
Mang	69	0	0	0	8	0	0	0	0	1
Urb	1	100	0	0	8	0	0	0	0	0
Sw	0	0	82	0	0	0	1	0	1	0
Tmp	6	0	0	86	1	6	3	0	1	9
Perm	25	0	0	2	81	0	0	0	0	9
Woo	0	0	1	0	0	90	0	0	0	0
For	1	0	13	1	2	0	92	0	6	0
Sav	0	0	0	0	0		0	100	0	0
Open	0	0	4	11	0	3	4	0	92	0
Raph	0	0	0	0		0	0	0	0	81

F)	Mang	Urb	Sw	Tmp	Perm	Woo	For	Sav	Open	Raph
Mang	71	0	0	1	8	0	0	0	0	0
Urb	7	90	0	0	0	0	0	0	0	0
Sw	7	0	76	0	0	0	5	0	1	0
Tmp	0	5	2	76	19	0	4	0	0	15
Perm	4	0	0	1	54	0	0	0	0	16
Woo	0	0	4	8	0	98	0	5	0	0
For	0	5	2	0	0	0	79	0	7	0
Sav	0	0	0	1	0	0	0	95	0	0
Open	11	0	16	12	0	2	11	0	92	0
Raph	0	0	0	0	18	0	0	0	0	68

the impact of topography on the classification accuracy, the influence of the training sets, and the problem of validating the thematic results.

A. Low Resolution Contribution

It is expected that the use of low resolution features will provide a smoother classification of homogeneous areas. We did not attempt to reduce speckle using a speckle filter because we believe that image variation at 100 m is mainly due to the intrinsic forest texture (GRFM data has 59 equivalent number of looks). Filters may improve classification and will be explored in future work. Nonetheless, we chose to use the available data in its original form. Like texture measures, low resolution approximations provide information on the spatial context of a pixel. It will be valuable for classification of classes whose discrimination is mainly given by radiometry. However, if the decision tree chooses not to use high resolution features, the spatial resolution of the output classification will be coarser.

The overall classification rate improves using LWLR and HWLR. This is observed comparing T2 (30% misclassification) with T4 (21%), and T5 (20%) with T7 (17%). From analysis of Table IV(C) nad (D), one can infer that the main contribution of using low resolution data is to differentiate the Open Forest from the Forest class. The classification rate of Open Forest increases from 19% to 80% (T2 and T4). The improvement results from the splits being chosen on the low resolution data where the histogram distributions are thinner. The same observation holds when comparing T5 and T7, where the separation between the two forests is improved in the latter (69–85%).

However, the use of splits on the low resolution data increases commission errors of Permanently Flooded vegetation class to the Low Mangroves class. The misclassification is observable throughout the maps of Fig. 4. Mangroves only grows near the coast where there is salt water. However, the image filamentary structure of the Low Mangroves is very similar to flooded filaments often seen along small rivers. This problem could be partly solved by using geographical information. However, in our case, only the SAR data is used. Thus, Low Mangroves that are far from the coast on the classification map should be interpreted as Permanently flooded vegetation.

Moreover, Table IV(D) and (A) show that there is a slight commission error increase of the Woody savanna into the Grass savanna class by 9%. This is most probably due to mixed pixels.

In T4 and T7, the final classification maps have a spatial resolution of 400 m. That is because no splits are chosen on the high resolution features.

B. Temporal Change

The two GRFM SAR acquisitions are snapshots in time. Thus, the detected temporal changes correspond to changes in the time interval of February to November 1996. Temporal changes should be reflected in the decision tree as a combination of high and low water data. There are two classes characterized by temporal changes: the Swamp (Sw), and Temporarily Flooded vegetation (Tmp).

Tmp is classified by a combination of low and high water amplitude values on all decision trees. The best results are obtained using LWLR and HWLR. A decisive split on the T4 and T7 at 132 on LWLR following a split at 124 on

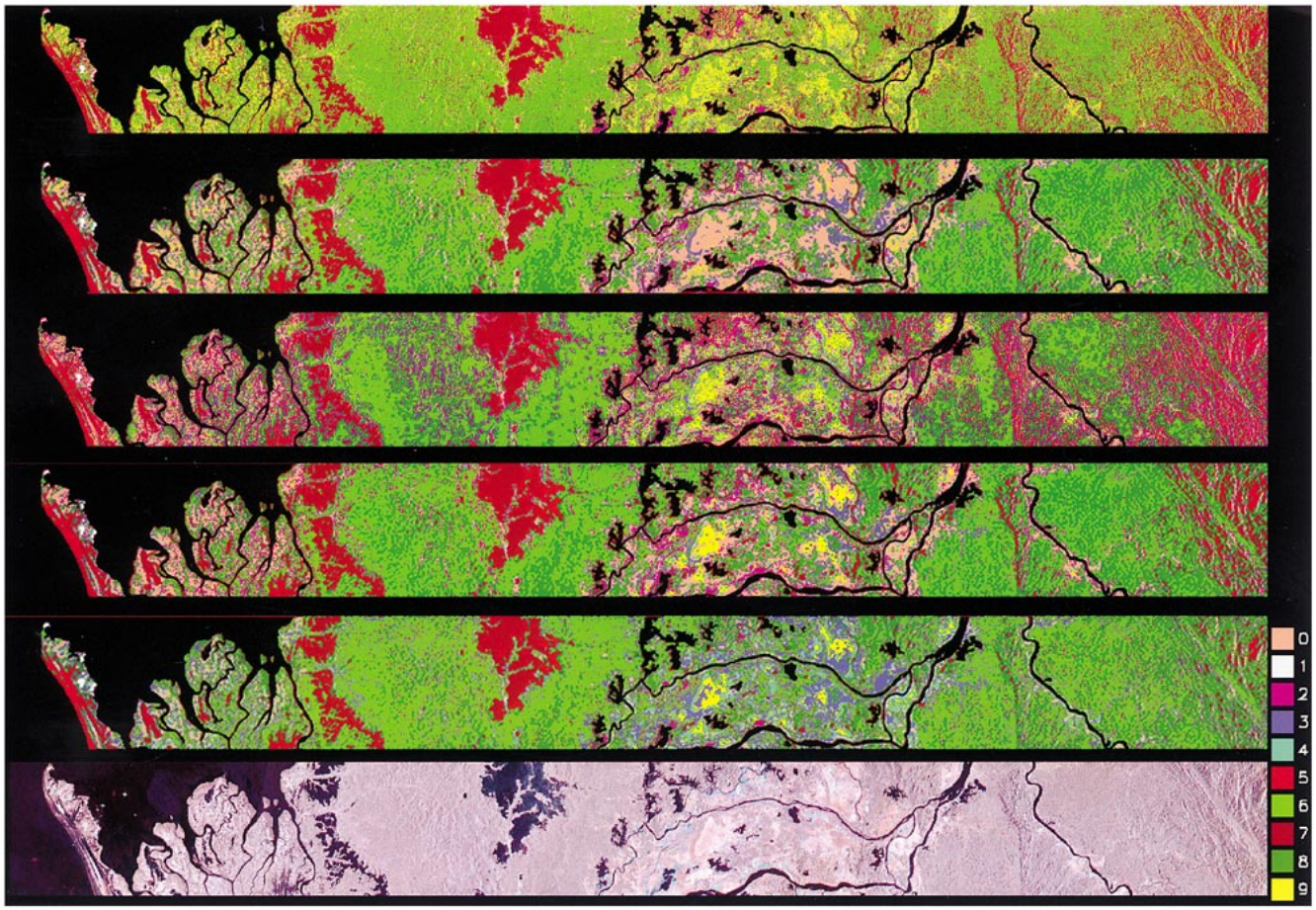


Fig. 4. Slices of the classified study area in Gabon. From top to bottom: Land cover classification obtained from decision trees T2, T4, T5, T7, T7I, and the red–blue composite of high and low water. The classes are Low Mangrove (0), Urb (1), Swamp (2), Permanently Flooded (3), Temporarly Flooded (4), Woody Savanna (5), Forest (6), Savanna (7), Open Forest (8), and Raphia (9), as defined in Table I.

TABLE IV
CONFUSION MATRICES FOR A) T7, B) T5, C) T2, D) T4, E) T7I, AND F) VALIDATION

class / nodesn	0 Mang	1 Urb	2 Sw	3 Tmp	4 Perm	5 Woo	6 For	7 Sav	8 Open	9 Raph
0	0	0	0	0	0	0	0	100	0	0
1	0	0	1	0	0	90	0	0	0	0
2	0	0	4	11	0	3	4	0	92	0
3	0	0	82	0	0	0	1	0	1	0
4	1	0	13	1	2	0	92	0	6	0
5	6	0	0	86	1	6	3	0	1	9
6	23	0	0	2	67	0	0	0	0	7
7	0	0	0	0	0	0	0	0	0	81
8	0	0	0	0	7	0	0	0	0	2
9	69	0	0	0	8	0	0	0	0	1
10	2	0	0	0	7	0	0	0	0	0
11	1	4	0	0	0	0	0	0	0	0
12	0	96	0	0	8	0	0	0	0	0

HWLR allows for distinction of 77% of Tmp (see Fig. 3). On the T2 and T5, the classification rates are only of 66% and 63%. Most of the misclassification occurs with Forest because of histogram overlap.

The Swamp class is also classified against other flooded classes by a combination of low and high water data. However, the high backscatter occurs in the low water image, when vegetation emerges from the water. The best results are obtained using T4 (91%) and T7 (91%). This is mostly due to the smoother estimation of the local amplitude provided by LWLR and HWLR.

Swamp is separated from Tmp by the opposite temporal trend. The opposite splits in T7 at 130_1 and 132_1 that follow 124_2 explicitly take into account their respective temporal trend (Fig. 3).

C. Texture Contribution

Decision trees T5 and T7 show that texture is used as a secondary feature because it is used late in the decision trees. However, texture contributes significantly to the discrimination of flooded vegetation classes such as Raphia, Permanently flooded vegetation and Low Mangroves. As can be observed in Table IV(C) and (D), those classes are poorly discriminated without texture. The main improvement is obtained for Raphia, which appears as smooth high amplitude regions in the radar images. As a general trend, Raphia is discriminated from Mangroves by using texture (for example on T7 mainly the 5_5 split with 82% $\pm$ 4% in terminal node 10).

Using texture in T5 also gave better discrimination of Permanently Flooded vegetation (69%) and Low Mangroves (80%). However, in T7, Low Mangroves and Permanently Flooded vegetation are poorly discriminated (node 11). The reason being

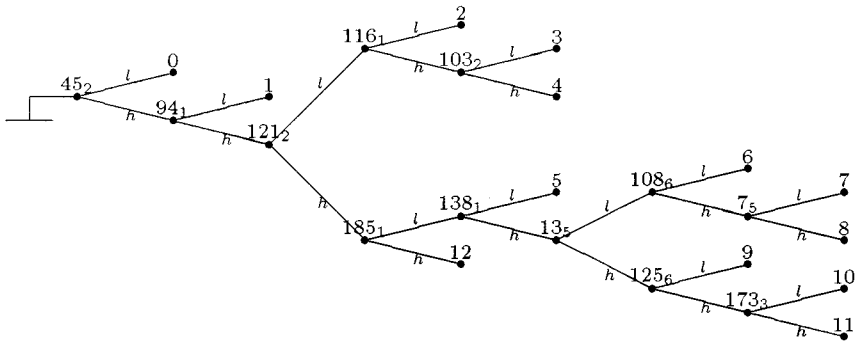


Fig. 5. Exploratory decision tree grown using testing training samples and pruning using growing samples with seven features. At each node, a value correponding to the best split s threshold is given with the input feature i (s_i). The features are the amplitudes for (0) low water, (1) low resolution low water, (2) low resolution high water, (3) high water, (4) texture 200 m, (5) 400 m, and (6) 800 m. The branches are denoted by l and h, which are lower and higher values than threshold s, respectively.

that early splits on the low resolution data have mixed samples of these classes.

From the previous discussions about feature contribution, we reach the following main conclusions.

1) The choice of input features is important when we consider classification rate of particular classes.
2) The overall best classification as searched by the Gini criterion is reached using all features as input.
3) The decision tree neglects information contained in unused features. This results in a lower spatial resolution classification (400 m).

To obtain different results, one could tune P_J in order to favor a given class j. Thus the Gini criterion would give more weight to purification of a class j. A second method would be to build many decision trees, which would each be specialized in classification of a given class. However, one needs to use a voting algorithm to decide on the final class of each pixels [29]. For this purpose, building of a large number of exploratory decision trees might improve classification.

D. The Topography

Topography is always an issue in remote sensing and particularly in SAR images (foreshortening and shadowing). Without a digital elevation model (DEM), it is not possible to retrieve the true land cover backscatter. In the Gabon image (Fig. 1), the east of the image is characterized by very strong topography induced radiometric and geometric effects. Mountain areas are mostly classified as Mangroves and Woody Savanna (see Fig. 4). In the first case, because of the bright filamentary geomorphology, which is very similar to Low Mangrove image pattern. In the latter case, the backscatter is similar to Woody savanna in either HWHR or LWHR. One of the effects of using low resolution data is to provide a smoother classification in the medium topography region located in the east–west center of the image.

E. Stability of the decision tree Structure

In this section, we discuss the issue of sensitivity of decision tree structure to training set variations induced by the choice of the training sites. The stability of decision tree structures was discussed by Breiman *et al.* [17, p. 155]. In fact, voting algorithms such as those discussed by Bauer and Kohavi [29] were introduced to deal with unstable classifiers like decision trees. However, in this paper, we do not attempt to use voting algorithms. Instead, we will test a single exploratory decision tree, built from the same independent training samples S_p and S_g. In order to build various training samples, one could have randomly subsampled S_p and S_g to obtain various correlated sample distributions or tuned P_j to build a large number of exploratory decision trees. For this experiment, we simply reverse the role of the training sets to build the exploratory decision tree. The exploratory decision tree will be grown with S_p and tested with S_g.

The competition between splits can be very tight when growing a decision tree. This often happens when features are correlated, which is the case for the amplitude images, both temporal and multiscale. Choosing a different split which had a similar cost (or reduction on impurity of the nodes) will lead to a different decision tree evolution [17, p. 156].

A seven feature decision tree (T7I) was grown using S_p, tested using S_g. The resulting decision tree is shown in Fig. 5 with the corresponding Table III. The land cover classification is shown in Fig. 6. Because the first split is different in T7I, the subsequent decision tree topology is also different but with a similar number of terminal nodes (T7 with 14 nodes and T7I with 13). The overall performance of T7I is better with 13% misclassification [see Table 4(E)]. Discrimination of Low Mangroves, Permanently flooded vegetation and Raphia improved significantly to 69%, 81% and 81% with lower commission errors. There is a slight increase of misclassification of Swamp into the Forest class.

Although the decision tree topology is different, there is consistency in the structure of the decision trees.

1) Temporal variation for Swamps is used for both decision trees for classification against Forests.
2) Similar splits (T7 at 124_2 and T7I at 121_2) distinguish Forests from the Flood classes (Mang, Tmp, Perm, and Raph).
3) Splits on TX800 classify Raph against Perm.
4) Splits on the TX400 classify Raph against Mang.

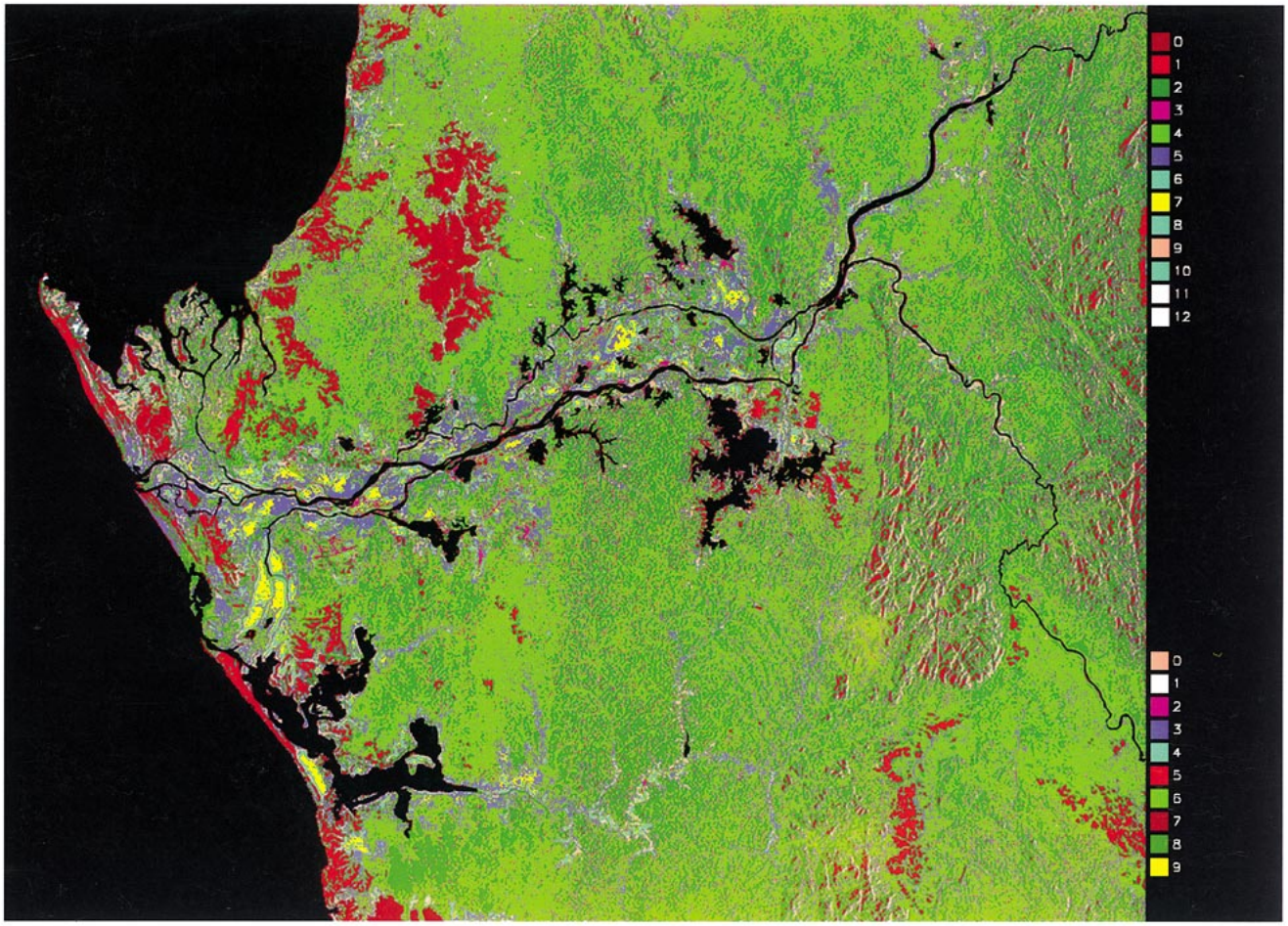


Fig. 6. Classified land cover map of the study area of Gabon using seven features with the exploratory decision tree. The 13 terminal nodes (top right) are labeled (in colors), with the class (bottom right) with largest membership. This land cover map is associated with the decision tree in Fig. 5 with Terminal node memberships as of Table III.

5) Combination of high and low water features discriminate Tmp against other flood classes.

F. Validation of decision tree Classification and Thematic Comparison

In the previous sections, we have used the decision tree structure to analyze the contribution and relations of features. We discussed decision tree performance based on the nominal misclassification, which is using S_g and S_p. To validate the decision tree classification,other authors have used a finite independent proportion of their training sample set to validate results [9]–[11]. Similarly, we selected a third independent training set S_v. The training samples were selected in different geographical areas than S_g and S_p for each class. The set S_v was input into T7I and the confusion matrix computed [Table IV(F)].

The overall decision tree classification accuracy is 80%, which is lower than the nominal overall accuracy. Most of the confusion occurs between the flooded classes such as Raph, Tmp, and Perm. The best classification results are obtained for classes with completely different backscatter characteristics, which are the savannas and Urban classes (i.e., very low and very bright backscatter). These classes were isolated early in the decision trees because of their distinctive backscatter. Forest is slightly mixed with Sw, Tmp, and Open forest.

VI. Summary and Conclusion

The use of a decision tree for classification of GRFM JERS-1 SAR images of tropical forest using multiscale information has been proposed. We chose ten classes which were in agreement with the vegetation maps provided by the "Ministère des eaux et forêts et du reboisement." The classes were: Low Mangrove, Urban, Swamp, Temporarily Flooded vegetation, Permanently Flooded vegetation, Woody Savanna, Forest, Grass savanna, Open Forest, and Raphia. Three spatially independent training sample sets were selected for each class. Those sets were used to grow the initial decision tree, test the pruned sequence and validate classification. In order to assess the contribution of the features, we built four decision trees using various feature combinations.

The main conclusions reached from the dimensionality analysis of the decision tree are the following.

1) Temporal changes are explicit in the decision tree structure.

2) Large scale texture measures are important features for distinction of flooded vegetation classes such as Permanently Flooded vegetation, Low Mangroves, and Raphia.
3) Categories such as savannas, forests, and flooded vegetation are discriminated by their amplitude backscatter values.
4) The decision tree classifier entirely neglects information contained in unused features. As a consequence in our case, the spatial resolution of the classification map is degraded. A tradeoff between classification accuracy and spatial resolution must be reached.

We also investigated the stability of the decision tree structure to training sets. Although a change of the decision tree topology occurred, the main structures and feature relationships are conserved. For example, temporal changes in Swamps and Temporarily Flooded vegetation are observed in both the seven feature decision trees. Moreover, texture refines the discrimination of flooded classes. However, performance of the decision trees are slightly different, a fact that suggests that exploratory decision trees using different training sets should be built.

The decision tree classifier allows the user to interpret each decision rule and make efficient use of only important features for the classification. These advantages will be valuable for thematic studies based on the GRFM JERS-1 mosaics.

ACKNOWLEDGMENT

The authors would like to thank Dr. L. White of the Wildlife Conservation Society, and Dr. C. Wilks, both in Gabon, for their invaluable guidance and information on the status of the vegetation cover in Gabon. They also wish to thank the staff of the Global Vegetation Monitoring Unit, Joint Research Centre, Ispra, Italy,for their outstanding work in the compilation of the GRFM Central Africa mosaic. The authors would also like to acknowledge the fundamental role of the Japanese Space Agency (NASDA), Tokyo, Japan, in the conception, development, and support of the GRFM project, which can be considered a landmark contribution to the global monitoring of tropical forest using radar remote sensing. Finally, they thank the anonymous reviewers for their help, useful comments, and contribution to the paper's improvement.

REFERENCES

[1] G. DeGrandi, J. Malingreau, and M. Leysen, "The ERS-1 Central Africa mosaic: A new perspective in radar remote sensing for the global monitoring of vegetation," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 1730–1746, May 1999.

[2] S. Saatchi, J. Soares, and D. Alves, "Mapping deforestation and land use in Amazon Rainforest by using SIR-c imagery," *Remote Sens. Environ.*, vol. 59, no. 2, pp. 191–202, 1997.

[3] E. Rignot, R. Zimmerman, and J. V. Zyl, "Space-borne applications of p-band imaging radars for measuring forest biomass," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, p. 1162, Sept. 1995.

[4] A. Luckman, J. Baker, T. Kuplich, C. Yanasse, and A. Frery, "A study of the relationship between radar backscatter and regenerating tropical biomass for spaceborne SAR instruments," *Remote Sens. Environ.*, vol. 60, no. 1, pp. 1–13, 1997.

[5] E. Nezry, E. Mougin, A. Lopes, J. Gastellu-Etchegorry, and Y. Laumonier, "Tropical mapping with combined visible and SAR spaceborne data," *Int. J. Remote Sensing*, vol. 14, no. 11, pp. 2165–2184, 1993.

[6] H. Kux, F. Ahern, R. Raney, R. Pietsch, and B. Tittley, "The contribution of SAREX'92 (South American radar experiment) campaign to the evaluation of natural resources in tropical rain forests: First results from test site acre, SW-Amazonia, Brazil," in *16th Symp. Canadien Sur la téelédétection and 8th congrès de l'Asociation Québécoise de télédétection*, 1993, pp. 53–59.

[7] A. Lopes, E. Nezry, R. Touzi, and H. Laur, "Structure detection and statistical adaptive speckle filtering in SAR images," *Int. J. Remote Sensing*, vol. 14, no. 9, pp. 1735–1758, 1993.

[8] M. Simard, F. DeGrandi, and K. Thomson, "Adaptation of the wavelet transform for the construction of multiscale texture maps of SAR images," *Can. J. Remote Sensing*, vol. 24, pp. 264–285, Sept. 1998.

[9] M. Friedl and C. Brodley, "decision tree classification of land cover from remotely sensed data," *Remote Sens. Environ.*, vol. 61, no. 3, pp. 399–409, 1997.

[10] R. D. Fries, M. Hansen, J. Townshend, and R. Sohlberg, "Global land cover classification at 8 km spatial resolution: The use of data derived from landsat imagery in decision tree classifiers," *Int. J. Remote Sensing*, vol. 19, no. 16, pp. 3141–3168, 1998.

[11] M. Hansen, R. Dubayah, and R. DeFries, "Classification trees: An alternative to traditional land cover classifiers," *Int. J. Remote Sensing*, vol. 17, no. 5, pp. 1075–1081, 1996.

[12] A. Belward and A. Dehoyos, "A comparison of supervised maximum-likelihood and decision tree classification for crop cover estimation from multitemporal landsat mss data," *Int. J. Remote Sensing*, vol. 8, no. 2, pp. 229–235, 1987.

[13] G. De Grandi, P. Mayaux, Y. Rauste, A. Rosenqvist, M. Simard, and S. Saatchi, "The global rain forest mapping project JERS-1 radar mosaic of tropical Africa: Development and product characterization aspects," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2218–2233, Sept. 2000.

[14] P. Mayaux, T. Richards, and E. Janodet, "A vegetation map of Central Africa derived from satellite imagery," *J. Biogeogr.*, vol. 26, pp. 353–366, 1999.

[15] A. Rosenqvist, V. Taylor, B. Chapman, M. Shimada, A. Freeman, F. D. Grandi, S. Saatchi, and Y. Rauste, "The global rain forest mapping project: A review," *Int. J. Remote Sensing, Special Issue on Global and Regional Land Cover Characterization from Satellite Data*, to be published.

[16] Y. Rauste, G. D. Grandi, T. Richards, A. Rosenqvist, G. Pena, E. Franchino, F. Holecz, and P. Pasquali, "Compilation of bi-temporal JERS SAR mosaic over the African Rain Forest Belt in the GRFM project," in *Proc. IGARSS'99* , 1999.

[17] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. New York: Chapman & Hall, 1984.

[18] J. Michaelson, D. Schimel, M. Friedl, F. Davis, and R. Dubayah, "Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys," *J. Veg. Sci.*, vol. 5, pp. 673–696, 1994.

[19] F. Ulaby, F. Kouyate, B. Brisco, and T. Williams, "Textural information in SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. 24, pp. 235–245, Mar. 1986.

[20] S. Mallat, "A theory for multi-resolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.

[21] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Processing*, vol. 4, pp. 1549–1560, Nov. 1995.

[22] E. Salari and Z. Ling, "Texture segmentation using hierarchical wavelet decomposition," *Pattern Recognit.*, vol. 28, no. 12, pp. 1819–1824, 1995.

[23] R. Fau, G. Bénié, J.-M. Boucher, and D.-C. He, "Segmentation Markovienne pyramidale d'images," *J. Can. Tele.*, vol. 20, no. 2, pp. 150–155, 1994.

[24] C. Lu, P. Chung, and C. Chen, "Unsupervised texture segmentation via wavelet transform," *Pattern Recognit.*, vol. 30, no. 5, pp. 729–742, 1997.

[25] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 710–730, July 1992.

[26] E. S. Kasischke, J. M. Melack, and M. C. Dobson, "The use of imaging radars for ecological applications—A review," *Remote Sens. Environ.*, vol. 59, pp. 141–156, 1997.

[27] L. Hess, J. Melack, and D. Simonett, "Radar detection of flooding beneath the forest canopy: A review," *Int. J. Remote Sensing*, vol. 11, no. 7, pp. 1313–1325, 1990.

[28] M. de F, M. Costa, E. de Moraes Novo, F. Ahern, F. M. II, J. Mantovani, M. Ballester, and R. Pietsch, "The Amazon floodplain through radar eyes: Lago grande de monte alegre case study," *Can. J. Remote Sensing*, vol. 24, no. 4, pp. 339–349, 1998.

[29] E. Bauer, "An empirical comparison of voting classification algorithms: Bagging, boosting, and variants," *Mach. Learn.*, vol. 1, pp. 1–38, 1998.

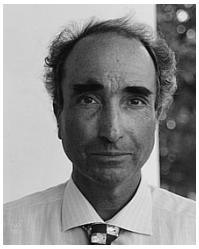
Marc Simard received the B.Sc. degree in physics and in astrophysics from the With Honor Program, Queen's University, Kingston, ON, Canada, in 1992, and the M.Sc. and Ph.D. degrees in physics (specializing in radioastrophysics) and geomatic sciences, respectively, from the Universite Laval, Quebec, PQ, Canada, in 1994 and 1998, respectively.

From 1995 to 1997, he was a Grant Holder with the Space Applications Institute of the Joint Research Center, European Communities (JRC SAI), Ispra, Italy. Since 1998, he has been a National Research Council Research Associate with the Jet Propulsion Laboratory (JPL), Pasadena, CA. His research interests include the multiscale analysis and classification of SAR images.

M. Simard is a U.S. National Research Council Research Associateship Awardee hosted at JPL.

Sasan S. Saatchi received the B.S. and M.S. degrees in electrical engineering from the University of Illinois, Champaign, in 1981 and 1983 respectively, and the Ph.D. degree from George Washington University, Washington, DC, in 1988 with a concentration in electrophysics and modeling of wave propagation in natural media.

From 1989 to 1991, he was a Postdoctoral Fellow, National Research Council, and worked at the Laboratory for Terrestrial Physics, NASA/Goddard Space Flight Center, Greenbelt, MD, on the hydrological application of active and passive microwave remote sensing. Since April of 1991, he has been a Scientist with the Radar Science and Engineering Section of the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, where he is involved in developing microwave scattering and emission models for soil and vegetated surfaces and retrieval algorithms for estimating geophysical parameters from spaceborne remote sensing instruments. He has been a Principal or Co-Investigator in several interdisciplinary international projects such as FIFE, EFEDA, Magellan, Mac-Hydro, Hapex-Sahel, BOREAS, LCLUC, and LBA. His present research activities include land cover classification, biomass and soil moisture estimation in boreal forests, land use and land cover change, and forest regeneration monitoring over tropical rain forests. His research interests also include wave propagation in disordered/random media and EM scattering theory.



Gianfranco (Frank) De Grandi (M'90–SM'96) received the Ph.D. in physics engineering (with honors) from the Politecnico Milano, Milano, Italy, in 1973.

Since 1977, he has been with the European Commission Joint Research Centre (JRC), Ispra, Italy, where he has performed research in signal processing for application areas such as gamma ray spectroscopy, data communications, and radar remote sensing. In 1985, he was a Visiting Scientist with Bell Communications Research, Morristown, NJ, where he participated in the design of METROCORE, one of the first research projects for Gb rate metropolitan area networks. From 1986 to 1989, he headed the signal processing section of the Electronics Division, JRC, where he introduced VLSI design technology and conducted research, in cooperation with Bellcore, on packet video, and in cooperation with ITALTEL Italy on the European digital mobile phone network. In 1989, he joined the Institute for Remote Sensing Applications [now Space Applications Institute (SAI)], where he started a research activity in radar polarimetry in the Advanced Techniques unit. Since 1994, he has been Principal Scientist for radar remote sensing in the Global Vegetation Monitoring unit. Since 1997, he has been Assistant Professor with the Faculte' de Foresterie et Geomatique, Universite' Laval, Quebec, PQ, Canada. His current research interests span a wide gamut, including global scale forest mapping using high resolution spaceborne SAR, multiresolution analyses based on the wavelet representation for texture measures, backscattering multitemporal estimators, topography sensing using polarimetric SAR data, and the statistics of polarimetric synthesized SAR images.

Dr. De Grandi is a senior member of the IEEE Geoscience and Remote Sensing and Signal Processing Societies and a member of the Planetary Society, Pasadena, CA.